

Evaluation of a theory of instructional sequences for physics instruction

Wackermann, Rainer; Trendel, Georg; Fischer, Hans E.

Postprint / Postprint

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

www.peerproject.eu

Empfohlene Zitierung / Suggested Citation:

Wackermann, R., Trendel, G., & Fischer, H. E. (2010). Evaluation of a theory of instructional sequences for physics instruction. *International Journal of Science Education*, 32(7), 963-985. <https://doi.org/10.1080/09500690902984792>

Nutzungsbedingungen:

Dieser Text wird unter dem "PEER Licence Agreement zur Verfügung" gestellt. Nähere Auskünfte zum PEER-Projekt finden Sie hier: <http://www.peerproject.eu> Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

gesis
Leibniz-Institut
für Sozialwissenschaften

Terms of use:

This document is made available under the "PEER Licence Agreement". For more Information regarding the PEER-project see: <http://www.peerproject.eu> This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Mitglied der

Leibniz-Gemeinschaft



Evaluation of a theory of instructional sequences for physics instruction

Journal:	<i>International Journal of Science Education</i>
Manuscript ID:	TSED-2008-0388.R1
Manuscript Type:	Research Paper
Keywords:	physics education, teacher development, experimental study, in-service, learning
Keywords (user):	physics education, teacher development, basis models



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Evaluation of a theory of instructional sequences for physics instruction

Abstract

The background of the study is the theory of *basis models of teaching and learning*, a comprehensive set of models of learning processes which includes for example learning through experience and problem solving. The combined use of different models of learning processes has not been fully investigated and it is frequently not clear under what circumstances a particular model should be used by teachers. In contrast, the theory under investigation here gives guidelines for choosing a particular model and provides instructional sequences for each model.

The aim is to investigate the implementation of the theory applied to physics instruction and to show if possible effects for the students may be attributed to the use of the theory. Therefore, a theory-oriented education program for 18 physics teachers was developed and implemented in the 2005/06 school-year. The main features of the intervention consisted of coaching physics lessons and video analysis according to the theory. The study follows a pre-treatment-post design with non-equivalent control group. Findings of repeated-measures ANOVAs show large effects for teachers' subjective beliefs, large effects for classroom actions and small to medium effects for student outcomes such as perceived instructional quality and student emotions. The teachers/classes that applied the theory especially well according to video analysis showed the larger effects.

The results showed that differentiating between different models of learning processes improves physics instruction. Effects can be followed through to student outcomes. The education program effect was clearer for classroom actions and students' outcomes than for teachers' beliefs.

Introduction

The background of this study is the theory of *basis models of teaching and learning* by Oser and Baeriswyl (2001) and its intended effects on physics instruction. The aim is to investigate the implementation of this theory and to show if possible effects for the students may be attributed to its use. The study follows a general framework for empirical investigations on teaching and learning by Fischer, Klemm, Leutner, Sumfleth, Tiemann & Wirth (2005).

The theory is a comprehensive set of models of learning processes, which include learning through experience, concept building or problem solving. It combines these models by giving guidelines when to choose which model in order to facilitate teaching and learning and also provides instructional sequences for each model. However, current physics instruction practice appears not to follow the theoretical requirements recommended according to the related literature (e.g. Reyer, 2004; Seidel & Prenzel, 2004). Therefore, a dedicated teacher education program is first necessary in order to create such theory-oriented instruction. A subsequent evaluation of that teacher education program can then possibly answer the question of the value of this theory. The overall aim is to **develop and validate a teaching strategy to facilitate learning of physics in schools.**

In this background section we will first describe the nature of the theory of *basis models* and then demonstrate why the theory seems promising to **function as teaching strategy in physics lessons**. Subsequently, we review the current state of teacher education programs before developing and presenting our own evaluation model guided by two research questions and detailed hypotheses.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

The theory of basis models of teaching and learning

There are many established teaching models in science education, which include discovery learning, scientific inquiry (Bybee, 2004; Lederman, 2004) and conceptual change (Posner, Strike, Hewson & Gertzog, 1982; Duit & Treagust, 2003) and articles and books have been written to provide summaries and overviews of different teaching models (e.g. Joyce & Weil, 1995; Treagust, 2007). Many of these models have been tested in comparison to more traditional teaching (e.g. Guzzetti, Snyder, Glass & Gamas, 1993; Lederman, 2007). However, it seems the combined use of different teaching models has not been fully investigated and it is frequently not clear under what circumstances a particular teaching model should be chosen.

In contrast, the theory of *basis models of teaching and learning* by Oser and Baeriswyl (2001) holds twelve different models of learning processes, each individually accepted in general educational theory, and gives guidelines when to choose which model according to teaching aims and learner’s prerequisites. Once a model is chosen, the theory gives guidance in planning and pursuing instruction by stating an instructional sequence.

The theory of *basis models of teaching and learning* distinguishes between surface structure and deep structure of instruction. The level of surface structure denotes all apparent aspects of instruction, which include methods such as individual learning or whole class instruction, demonstrations or lab work etc. (Oser, 2001; Reyer, 2004). The level of deep structure describes the teacher’s intentions and cognitive processes of the students while they proceed in their learning process. Thus, the proposed instructional sequences are related to intended learning processes of the students and deal with the deep structure of instruction.

Next, the theory differentiates between different instructional aims and assigns to each aim a specific model of learning processes. Each of these models of learning processes is one *basis model of teaching and learning*. Oser and Baeriswyl (2001) claim that these models have substantially different teaching aims and require substantially different prerequisites of the students. The models are operationalized as chains of lesson phases thus forming instructional sequences. Once a model is chosen, learners have to work in the right order through all of its phases in order to enable or facilitate learning. For Oser and Baeriswyl, instruction on the deep structure level is rather strictly ruled, whereas teachers have great freedom on the surface structure level. For example, if students are supposed to solve a problem, then whether they do group work or work individually is not so important, but first every student should develop a clear problem understanding, secondly propose solution attempts, subsequently put them to a test and so on. In contrast, if students are supposed to learn a new concept, completely different steps are necessary: First, rehearsal of existing pre-knowledge, then introduction of a prototype representing the new knowledge, then clarification of the new knowledge in this prototype and so on. In this way, each teaching aim has its own instructional sequence. Oser and Baeriswyl (2001) call this *Choreographies of teaching*, where like in a dance the underlying rhythm is fixed and dependent on the dance (deep structure, dependent on learners' needs and teaching goals), whereas individual styling is more flexible (surface structure, e.g. group work or not).

Why does the theory of basis models of teaching and learning seem promising to improve physics instruction?

In brief, physics as a school subject across Western countries is not well received by students and student performance is often below expectations (Goodrum, Hackling &

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Rennie, 2001; Osborne & Dillon, 2008). Studies and reports often draw on results of large-scale assessments like TIMSS, PISA or the ROSE study. Although neither of these studies allows statements about the instruction, it can be assumed that student performance and attitudes are dependent on instruction.

Results from the TIMSS-video-study in mathematics from the USA and other countries (Pauli & Reusser, 2003; Stigler & Hiebert, 1997) indicate that variances in student performance could not be explained by aspects of instruction involving surface structure. Instead, it is believed that the deep structure of instruction has an effect on student learning (Hofstein & Lunetta, 2004; Klieme, Lipowsky, Rakoczy & Ratzka, 2006; Reyer, Trendel & Fischer, 2004, Seidel, Prenzel, Rimmele, Schwindt, Kobarg, Herweg & Dalehefte, 2006). In physics instruction Seidel et. al (2006) have successfully demonstrated that aspects of deep structure like systematic experimenting have positive effects on students' attitudes and academic achievement.

Other attempts to improve physics instruction aim at an improved surface structure of instruction, like the inquiry approach in the USA (Bybee, 2004; Lederman, 2004) or the problem-oriented approach in Germany (Fries & Rosenberger, 1970; Schmidkunz & Lindemann 1992; Wagenschein, 1970). Though teaching models like *inquiry* are promoted for more than two decades, overall instructional results are not seen as satisfactory (Lederman, 2004). For other models as the sole focus of learning process in science instruction, the results seem more promising at first, such as in the case of *conceptual change* reported in the meta-analysis by Guzzetti et al. (1993). However, study results were often not achieved with multiple sources of evidence, such as interviews and classroom observations, which is needed to judge conceptual change (Duit & Treagust, in press). In fact, there is now some

1
2
3 experimental evidence that for example the inquiry approach is not better than
4 traditional teaching (Lederman, 2007). A possible cause might be an inadequate use
5 of these teaching models and their instructional sequences. Maybe, teaching models
6 are used inappropriately with respect to learners' prerequisites or phases of their
7 instructional sequences are omitted. Labudde and Duit (2007) for example report
8 from a Swiss-German comparative study that German physics instruction is most of
9 the time structured according to only one single learning process, the problem-
10 oriented approach (Kircher, 2007). Promoting only one learning process neglects the
11 fundamental philosophical works by Popper, Kuhn and others, who showed that
12 there are many ways to reach new insight in science using different scientific
13 methods. Experiences, for example, do not directly lead to scientific concepts,
14 because concepts usually are not discovered by individuals but are created in a
15 complex social process (Lederman, 2004).
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32

33
34 The theory of *basis models of teaching and learning* may offer a remedy by providing
35 a variety of instructional sequences with rules for which model to choose according to
36 the learners' needs and the teaching goals. In other words, the theory gives guidance
37 for the deep structure of instruction, where instruction is organized according to the
38 learners' needs, and at the same time promotes **a strategy to apply** a variety of
39 learning processes. The combination of scaffolding instruction according to the
40 learners' needs and allowing for different learning processes again paying attention
41 to the learners' needs seems promising to improve current physics instruction. The
42 following paragraphs describe the circumstances for the selection of a *basis model of*
43 *teaching and learning* when planning physics instruction.
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Different models of learning processes at work

According to Reyer (2004), the following learning processes are seen as most important for physics instruction: Problem solving, concept building and learning through experience. The instructional sequences with the lesson phases for these three models are shown in Table 1.

[Insert Table 1 about here]

The individual learning sequences can be traced back to several different roots. Problem solving has been best investigated and goes back to the work of Dewey (1933), Polya (1957), Bransford (1984) and Dörner and Wearing (1995). The above stated learning sequence is found to be identical in all their references. Learning through discovery as a subdomain of learning through experience has been described widely, for example by Dewey (1974), Bruner (1961) and Wagenschein (1970). The model of situated learning also aims for experiential learning (Lave & Wenger, 2003). Although the models are not fully congruent, it is stressed that personally meaningful learning should involve students' own experiences. The epistemological roots of concept building trace to Aebli (1981, 1983). Often cited together with the model of conceptual change (Posner, Strike, Hewson & Gertzog, 1982), which is not quite the same, concept building builds on connecting and hierarching existing knowledge into higher concepts thus forming a web-like knowledge system.

The learning goals or outcomes of these three learning sequences are completely different. This can for example be seen in the final lesson phases, where learning through experience leads to a general rule like Ohm's Law, concept building leads to the development of a theory or a model like conservation of energy or momentum,

and problem solving leads to a solution strategy, not to factual knowledge, such as how to tackle problems in mechanics involving time dependency.

The prerequisites for students to successfully follow these sequences are also different. For learning through experience prior knowledge is less crucial than for problem solving. The latter only works when all necessary knowledge is already available to the learners, since no factual knowledge is gained during the process. Concept building builds on prior students' knowledge but students are guided towards new insight. To clarify, research certainly indicates that students can/will interpret experiences in terms of prior knowledge, but learning through experience does not require prior knowledge in the same way as, for example, problem solving does.

In summary, the theory of *basis models of teaching and learning* is based on ample research evidence in psychology and pedagogy. It combines teaching and learning by differentiating between characteristic learning goals and addressing to each of these specific learning processes, which are operationalized as chains of lesson phases. The appropriate use of these *basis models*, dependent on teaching aims and learners' prerequisites, seems promising to facilitate learning and is the focus of the article.

Modifications to the theory

The theory was modified and specified in our research group to meet the particular characteristics of planning and analysing physics instruction (Fischer, Trendel, Reyer & Wackermann, 2006; Trendel, Wackermann & Fischer, 2007). Firstly, we translated Oser's original description into the language of physics instruction. Secondly, we added so-called levels of fulfilment. Oser's original description of the individual lesson phases describes an ideal way of instruction, which we did not find in our pilot study

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

and appears to be unrealistic for current physics instruction. Hence our levels of fulfilment gradually describe the full realisation of the theoretical requirements for each lesson phase. See Table 2 for the example of lesson phase 2 in the basis model of learning through experience. These levels were provided for the three most important teaching models, problem solving, concept building and learning through experience and for each instructional phase. Level 0 always denotes a superficial performance of the students and level 3 always denotes the complete fulfilment of an instructional phase. In summary, the levels describe to which degree the students perform according to the theory, regardless of the actual teaching model or instructional phase. The levels were also used for the teacher education program, for example when teachers' expectations and students' performance did not match. Examples include instances when teachers chose a teaching model which fitted not to the students' prior knowledge or when individual phases of a well-chosen instructional sequence were omitted and subsequently the students were able to fulfil the next phase's demands only at a very low level of fulfilment.

A recent work by Kirschner, Sweller and Clark (2006) shows that strong guidance is important for novice students when doing inquiry. Following this finding and the original work of Oser and Baeriswyl (2001), the levels of fulfilment for learning through experience reflect more and more internal guidance.

[Insert Table 2 about here]

Research on evaluation of teacher education programs

Teacher education programs are abundant but so far were seldomly evaluated (Garet, Porter, Desimore, Birman & Yoon, 2001; Terhart, 2003). Only a few studies investigated improvements in classroom actions or student performance as an outcome of teacher education (Fishman, Marx, Best & Tal, 2003). A good overview is

given by Lipowsky (2004) in a meta-review. In accordance with the work of Fischler (2003) evaluations of learning-process-oriented teacher education programs in the sciences are seldom reported and show small effects.

A study by Hand and Treagust (1994) designed to alter the subjective beliefs of eight volunteer teachers over the course of more than one year towards a constructivist view on teaching and learning proved to be successful, but it remained unclear whether the teachers' classroom actions changed as well. Effects on students were not investigated. The authors group their work as a qualitative study, since data was acquired by means of qualitative interviews.

For more than one year, Luft (2001) provided in-service education to a group of 14 secondary school teachers with respect to inquiry learning. Six of the 14 teachers were young professionals, the others were experienced teachers. By means of partially category-based interviews and classroom observation data teachers' subjective beliefs and classroom actions were identified. In the case of the experienced teachers, no changes in subjective beliefs were reported, but aspects of intended classroom actions were observable. In the case of the young professionals, the subjective beliefs changed, but not their classroom actions. Effects on students were not investigated.

A study by Yerrick, Parke and Nugent (1997) also aimed for altering the subjective beliefs of in-service teachers towards a more constructivist view of teaching and learning. Eight volunteer teachers were **interviewed** by means of semi-standardised interviews before and after a two-week summer school. Results of the study show that the teachers merely adapted those views. Effects on classroom actions or on students were not investigated.

Mikelskis-Seifert and Duit (2007) studied the effects of the German teacher education program called “Physics in context”. The main goal of the intervention was to change teachers’ subjective beliefs about teaching and learning in order to improve instructional quality. Results of the evaluation comprise effects on student interest and self concept, but classroom actions were not controlled.

In summary, all the above mentioned studies were designed to improve learning for the students by altering teachers’ subjective beliefs. The methodologies for evaluation were mixed, with teachers’ subjective beliefs always being investigated, but classroom actions or student effects only sometimes, and never all together.

Our model for evaluation of the teacher education program

We will now present our model for the evaluation of a teacher education program. **As confirmed in many studies** teachers’ subjective beliefs, classroom actions and student outcomes are equally important **for assessing the quality of a teaching process. Thus, we** follow an early work by Kirckpatrick (1979). Our model comprises of teachers’ subjective beliefs, of classroom actions and of students’ outcomes – their perceptions, students’ emotions and students’ performance. Student perception denotes a wide range of classroom observations of instructional quality, which include clarity, structuredness and pacing. Students’ emotions include student interest, motivation and self concept. Our evaluation model is shown in Figure 1.

[Insert Figure 1 about here]

We investigate student effects because a learning-process-oriented teacher education program is designed to improve learning for the students. Simultaneous control of classroom actions and of teachers’ subjective beliefs allows the decision whether possible effects on students are really due to the intervention. The indicated arrows represent how we expected the intervention to work. Teachers’ subjective

beliefs act on classroom actions according to the model of “Subjective Theories” developed by Groeben, Wahl, Schlee and Scheele (1988). Changes in teachers’ subjective beliefs should lead to changes in classroom actions. The arrow from classroom actions to students’ perception, emotions and performance draws on a model of teacher’s offer, mediation and moderation and students’ use in instruction (Helmke, 2003). This model refers to a general instruction model of Helmke and Weinert (1997), in which teachers’ actions are regarded as an interaction between an offer to the students and their opportunity to learn.

Research questions

Consequently, we formulate two research questions: Is the education program effective on teachers’ subjective beliefs, on classroom actions and on student outcomes? If there are effects for the students, do they correlate with basis-model-like teaching?

We hypothesised for teachers’ subjective beliefs for example that the importance of student competencies increases. Competencies denote a wide range of student abilities such as “classical” problem solving, planning of investigations to receive new insights, applying thinking patterns and allowing critical analysis of results achieved. We hypothesised that because of the education program teachers would value such student competencies higher than for example factual knowledge.

With regard to classroom actions we expected that teachers and students increasingly follow the theory. For example, we expected an increased diversity of use of teaching models and increased levels of fulfilment.

With regard to the students we expected that theory-based instruction can be followed easier by the students thus improving student perception in areas which

include clarity and structuredness. Possibly there are effects also on student interest, motivation and performance.

With regard to the second research question we expected that classes following the theory especially well should show larger effects for the students.

Methods

Sample

The intervention group comprised of 18 teachers (16 male, 2 female) from nine different upper level secondary schools in Germany. All participating teachers were volunteers and therefore probably to begin with they were motivated to try something new. In the school year 2005/06 these 18 teachers took part in an education program directed towards the *basis models of teaching and learning*. Each teacher focused on one of his/her classes. The age level of students lies between 13 years and 18 years, with the majority between 15 and 16 years. The implemented education program used a theoretical model for altering teacher behaviour, which was published by Fischer, Trendel, Reyer & Wackermann (2006) and Trendel, Wackermann and Fischer (2007). The main features of the intervention comprised coaching of physics lessons, video analysis according to the theory of basis models, and post-reflection with teachers. The coach assisted the teachers in choosing a model of learning process suitable for the general teaching aim and the prerequisites of the learners. He also assisted in planning the instructional sequence and in planning individual instructional phases. Coaching objectives were to plan instruction according to the theory so that lesson objectives and lesson structure become clearer to students as well as highlighting the processes in physics. Processes again denote a wide range such as “classical” problem solving, planning of investigations to receive new

insights, applying thinking patterns and allowing critical analysis of results achieved. We planned that every participant went through this coaching cycle five times. It turned out, however, that some teachers were only able to do that less than five times whereas others did it six times, either because teachers became ill or because of changes within the schools or because they wanted more coaching. Parallel to the intervention group a comparison group of 17 teachers/classes was formed, mostly from parallel classes of the same school. Teachers of the control group did not take part in the education program and no videos were taken. They and their students only took part in the questioning before and after the intervention, **presumably** rendering their instruction unaltered from their normal routines. As stated earlier, current physics instruction in Germany commonly follows one teaching model only, the problem-oriented approach, which is dominated by times of teacher-centered direct instruction and times of student-centered lab work (Reyer, 2004; Fischer, Klemm, Leutner, Sumfleth, Tiemann & Wirth, 2005; Labudde & Duit, 2007).

Variables

For variables investigated, there is one independent variable, either intervention or comparison group, and a number of dependent variables.

For teachers' subjective beliefs, the education aims like the importance of student process competencies and professional aims like the importance of own professional development were sought from the participating teachers. The scales, four of them with 4 to 5 items each, were developed for this study and piloted. For classroom actions, three variables were investigated: the basis model, instructional phase and level of fulfilment. The first two are direct derivatives of the theory of *basis models of teaching and learning* by Oser and Baeriswyl, though operationalised for physics instruction. The last one is our own invention. For students' outcomes, eight variables

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

with 4 to 6 items each were investigated: Students' perceptions of lesson quality, which includes clarity and structuredness, and student emotions, which include interest, motivation and self-concept. The student variables originated in the TIMSS-background questionnaire in German (Baumert, Bos & Watermann, 2000) and in a lesson quality instrument developed in German by Clausen (2004). The student performance was tested with a sample of TIMSS items from population II and population III (TIMMS-95; TIMSS-99). In addition, the student cognitive abilities were controlled by means of a short version of a standardised cognitive ability test (Heller & Perleth, 2000).

The reliability of all questionnaire scales lies between $0.6 < \text{Cronbach's alpha} < 0.9$ for both pre and post assessments and, where applicable, resemble those values found in literature. For the video analysis, 13 randomly chosen videos, totalling 513 minutes of instruction time, were analysed by two different coders. With a total of 80 videos, this is more than the commonly agreed ten percent of double coding. The interrater-agreement turned out to be Cohen's kappa = 0.8 for the coding of both basis model and instructional phase and Gutmann's gamma = 0.7 for levels of fulfilment.

A selection of variables used in this study is listed in table 3 along with sample items, reliability and origin. For results and discussion, not only those listed here but all variables under investigation were considered.

[insert Table 3 about here]

Study design and data acquisition

The study follows a pre-post design (Cohen, Manion & Morrison, 2000) to compensate for possible differences between comparison group and intervention group before the start of the intervention. The study design is quasi-experimental

(Cohen et al., 2000) because the intervention group is not randomized. Pre-assessments took place in September/October of 2005 and post-assessments in June 2006, respectively. Teachers' subjective belief data and student perception and emotion data were acquired by means of questionnaires. The classroom actions were investigated by means of a category-based video analysis with one-minute time intervals.

Planned statistical analyses

For research question 1 (Are there effects on teachers' subjective beliefs, on classroom actions and on student outcomes?) repeated-measures ANOVAs were performed checking for interaction effects between time and group for all student and teacher variables. In the case of classroom actions, no comparison group videos exist. Hence, analyses of variance are not possible and no causal relations can be found. Instead, simple descriptive analysis was used to detect changes over time in the classroom actions as the intervention progresses.

For research question 2 (Can possible positive effects for the students be correlated with basis-model-like instruction?) the above mentioned student investigations were repeated while controlling for classroom actions. This was done by separating the intervention group in high- and low-performing teachers/classes according to their theory-orientation and subsequent repetition of repeated-measures ANOVAs with three groups (high performer, low performer, comparison group). Theory-orientation may be measured for example with the average level of fulfilment.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Results

Teachers’ subjective beliefs

For teachers’ subjective beliefs there was an expected, significant and large interaction effect ($F(1,29)=3.87$, $p = 0.016$, $\eta^2 = 0.184$) between group and time for perceived importance of student competencies, as can be seen in Table 4. This effect is due to two simultaneous changes. Teachers of the intervention group report an increased importance of student competencies over time whereas teachers from the control group report a decreased importance over the same time. This decrease seems surprising at first but might be due to the fact that post-assessments took place towards the end of the school year when teachers not taking part in an education program might lower their expectations. Overall, this is a large positive effect for the intervention group.

[insert Table 4 about here]

Another expected interaction effect occurred for importance of professional development ($F(1,29)=5.218$, $p=0.030$, $\eta^2 =0.152$). As expected, there was no significant intervention effect for the importance of declarative knowledge. Also as expected for this type of intervention aiming at structuring physics instruction, there was no significant intervention effect for the importance of a positive attitude towards physics. Briefly summarized, the education program could significantly and as expected change the intervention group teachers’ views on the importance of the instructional aims. In addition, the perceived importance of professional development was raised for the intervention group teachers.

Classroom actions

With regards to classroom actions there were no control group videos. Hence, only changes over time within the intervention group can be reported. Changes can be identified from the first videotaped lesson of each teacher, which was taken before the intervention started, to the 62 videos taken during the education program. This way, the 18 first videos serve as a base-line. Variables under investigation comprise the basis model, instructional phase and level of fulfilment. Expected changes include a more diverse use of basis models, a more substantial use of individual instructional phases and an increase in level of fulfilment over time.

Before the intervention started, instruction followed with about equal use of the model of learning through experience and of concept building, as shown in Table 5. The model of problem solving played only a marginal role. During the coaching the model of problem solving is more pronounced. Hence, the variety of models of learning processes increased in the course of the education program.

[insert Table 5 about here]

In the education program, some instructional phases were especially stressed, because they seemed to play a key role in their respective instructional sequences. Among these phases was problem solving phase 2, the precise problem specification, and concept building phase 1, the rehearsal of the pre-existing knowledge. An investigation of changes from before the intervention to during the intervention yields the following results: The average length of phase 2 of problem solving seems to increase drastically from 2.5 minutes before the intervention started to about 8.2 minutes during the intervention. However, a paired t-test ($t(11)=1.71$, 2-tailed, $p=0.12$) proves not to be significant. This is probably due to the small sample size, i.e. there were only few lessons that actually followed the teaching model of

problem solving. For the length of phase 1 of concept building there is an increase from on average 3.00 minutes before the intervention to 5.25 minutes during the intervention. However again, a paired t-test (2-tailed) with $p=0.10$ and $t(55)=1.70$ proved not to be significant. Similar results were obtained for lesson phase 3 of **learning through experience**, the first reflection. In summary, there is no clear evidence that the education program increased the duration of individual instructional phases.

The levels of fulfilment undergo significant changes in the course of the intervention. The average time that students participated at level of fulfilment 0 decreased from on average 9.8 minutes per lesson before the intervention started to 7.0 minutes during the intervention. A paired t-test (2-tailed) with $p=0.03$ and $t(76)=2.30$ proved to be significant (effect size $d=0.5$). In contrast, the average time students participated at level of fulfilment 2 increased significantly from 3.3 minutes to 8.0 minutes (paired t-test, 2-tailed, $t(57)=2.31$, $p=0.02$, effect size $d=0.8$). So there is evidence from the levels of fulfilment that the intervention did change classroom actions. Level of fulfilment 2 turns out to be the highest participation level that students noteworthy reached in all 80 videotaped lessons. **The reader should know that the lessons planned by the teacher and the coach not always intended level 3, for example, when students were not yet used to work according to their own plan. In those cases level 2 was planned for.**

To elaborate the effect of the intervention in more detail, the intervention group will be divided into high (IG High) and low performing teachers (IG Low), taking into consideration the suggestion of the coach. According to his judgment, a subgroup of five to eight teachers can be separated. In order to make clear statements, in this article only the strongest five teachers are considered. They form the

group of high performing teachers (IG High), who followed the theory especially well. The remaining 13 teachers, who according to the coach's judgement lesser altered classroom instruction, form the low performing group (IG Low). Data from the video analysis supports this hypothesis.

[insert Figure 2 about here]

Figure 2 shows, separately for high (IG High) and low performer (IG Low), the average time students participate at level of fulfilment 2 in comparison between before and during the intervention. As one can see in Figure 2, students of both groups initially participated only marginally at level of fulfilment 2. An independent samples t-test (2-tailed) shows no significant difference between both groups. During the intervention time lengths increase and an independent t-Test for the difference between the groups becomes significant ($t(45)=3.0$, 2-tailed, $p<0.01$, effect size $d=1.0$). In addition, the increase from the first video to videos taken during the intervention proved to be significant only for the high performers (paired t-test, $t(17)=4.4$, 2-tailed, $p<0.01$, $d=1.5$), whereas the increase for the low performers is not significant.

In summary, there were significant and expected changes in the classroom actions. In addition, the intervention group can be split up in a high and a low performing group with classroom actions of the high performing group being significantly more theory-like than those of the low performing group. For the analysis of students' data three groups are considered: Classes of high (IG High) and low performing teachers (IG Low) and the comparison group (CG).

Students' perception of instructional quality

For students' perception of instructional quality the largest effect with $\eta^2 = 0.038$ occurs for the variable of perceived teaching for understanding ($F(2,722)=14.22$;

p<0.01). Figure 3 shows the interaction effect between time and group with all three groups being separate. Students of the comparison group report no change in their perception of teaching for understanding over time. In the same time, students of the low performing group (IG Low) show an increase of perceived teaching for understanding and students of the high performing group (IG High) show an even bigger increase. Pair-wise analyses of variance confirm the significance of these separations; the interaction effect between IG High and CG is mid-sized ($\eta^2 = 0.054$, $p < 0.001$).

[insert Figure 3 about here]

For other investigated variables of student perception of instructional quality, interaction effects occur only for students of the high performing group. In these variables, the low performing group cannot be distinguished from the control group. This situation holds true for perceived lesson pacing ($F(2; 722) = 10.25$, $p < 0.001$, $\eta^2 = 0.028$), and also for perceived structuring aids ($F(2,722)=5.47$, $p < 0.01$, $\eta^2 = 0.015$). Additionally, a dummy variable was investigated, i.e. a variable, in which no change for both groups should occur. By asking a dummy variable, one can show that only intended effects happened. In this case, students were asked for perceived individualisation of instruction. An exemplary item reads as “In our physics class more talented students may proceed faster”. Since the education program did not cover this aspect of teaching, no effect was to be expected and indeed all three student groups report almost identically over time.

On the level of student emotions there are significant effects only for students of the high performing group. Figure 4 shows as an example the variable interest in physics. There is an interaction effect ($F(2,722)=8.77$; $p < 0.01$) only for students from

the IG High. The pair-wise comparisons between IG High and IG Rest and between IG High and CG show mid-sized effects of $\eta^2 = 0.030$ and $\eta^2 = 0.039$, respectively.

{insert about here Figure 4}

Further investigated variables include student intrinsic motivation with an interaction effect again only for the IG High $F(2,722)=4.20$; $p=0.01$). The pair-wise comparisons between IG High and IG Low and between IG High and CG turn out to be equal in size, $\eta^2 = 0.018$, respectively. The postulated interaction effect for physics self-concept turns out to be only marginally significant $F(2,722)=2.74$; $p=0.07$). An additionally tested dummy variable asking for student extrinsic motivation yields no effect at all. Since the education program focused on elements of instructional sequencing, but not on schooling conditions, no effect was to be expected for students' reports on extrinsic motivation.

An additionally administered TIMSS-performance test supports the findings above. A t-test shows mean score differences between IG High and IG Low ($t(388)=2.53$; $p=0.012$) as well a t-test between IG High and CG ($t(446)=2.04$; $p=0.045$), each with an effect size of $d=0.2$. Although a one-time performance test cannot show causal relations, the results point in the same direction as the other findings in favour of the high performing intervention group.

In summary there are a number of expected intervention effects for classes from high performing teachers in areas of students' perceptions and students' emotions. Effect sizes are small to mid-sized. Students of low performing teachers report a difference compared to the students of the control group only for one variable of classroom perception. Results of a TIMSS performance test do not rule out that such theory-oriented instruction might impact student performance as well.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Teacher beliefs again

Subsequently, analyses of variance for the teachers’ subjective beliefs were repeated with three groups. In the area of educational aims, the beliefs change for the entire intervention group homogenously compared to teachers of the control group. So no new insight is gained here. However, for the one variable asking for importance of continuing professional improvement, there is an astonishing finding. The previously reported significant interaction effect is solely due to teachers of IG High, as one can see in Table 6. These teachers start much lower to begin with and then show a drastically increased importance of such professional aims to levels comparable to the other groups. A pair-wise comparison between IG High and CG proves to be significant ($F(1,17)=16.98$, $p<0.01$, $\eta^2 = 0.50$) and yields the largest effect found in this study at all.

[insert Table 6 about here]

Conclusions and limitations

The aim of this study was to investigate whether physics instruction planned and conducted according to the theory of *basis models of teaching and learning* yields benefits for the students. Therefore, a dedicated teacher education program was evaluated in terms of teachers’ subjective beliefs, classroom actions and student outcomes. The study followed a quasi-experimental pre-post design.

Results of teachers’ subjective beliefs show that the education program was able to change the educational and professional aims of the teachers of the intervention group mostly as hypothesised. Effect sizes are considered large according to Rost (2005).

Regarding classroom actions, large changes according to the theory of basis models can be seen. The variety of models of learning processes in use increased and the levels of fulfilment also rose. In addition, the intervention group can be split in two groups, high and low performers, each following the theory to a different degree.

Students especially from high performing teachers/classes report the hypothesised intervention effects, small and middle in size, in areas of perception of instructional quality and student emotions. Students from low performing teachers/ classes show only a small intervention effect for one variable of perception. These findings are additionally supported by results from a TIMSS performance test. Regarding teachers' subjective beliefs, however, high and low performing teachers can be distinguished by only one variable.

According to Fan (2001), effects with both statistical significance and at least medium size can be reported with confidence, which is the case for results in all investigated areas.

In conclusion, research question 1 ("Is the education program effective for teachers' subjective beliefs, classroom actions and students' outcomes?") can be answered positively, although effects on classroom actions and on student outcomes are different in size for classes of high and low performing teachers.

Research question 2 can be answered positively as well. The classes that followed the theory especially well according to video analysis show the biggest effects on student variables. It is through the analysis of videos that effects for the students can be traced back to changes in classroom instruction. In addition, dummy variables for teachers' beliefs and students' perception/emotions validate that only the intended effects of the education program occurred leaving other areas of teaching and learning untouched.

Concerning limitations of this study, it can be argued that the reported effects are due to the mere presence of a(n experienced) teaching coach. However, we controlled by means of video analysis whether the teachers actually followed the intended basis model paradigm. Teachers who followed the basis models more so were put for the analysis into IG High, the others were put for the analysis into IG Low. In some way, IG Low can be considered as the real control group, where teachers of both groups received identical coaching. This renders an experimental setting **with control for the involvement of the coach**. We kept the ordinary control group in the analyses, because it can be assumed from literature (see above) that instruction in the control group follows the problem-oriented approach. Videos from IG Low still resemble this traditional teaching approach using only one model of learning processes.

Additionally, an improvement of instruction due to the mere presence of a coach should affect all kinds of instructional quality. But the student investigation shows that only intended effects occurred and not general, arbitrary enhancements of instructional quality. We therefore conclude that instruction according to the basis models caused the reported effects, and not the mere presence of a coach.

Finally, it can be argued that the results are true only for a special subgroup of motivated, selected physics teachers. True, but those selected, motivated teachers who were able to implement the paradigm more (IG High) than the others (IG Low) also increased their (high to begin with) instructional quality more. We therefore **assume that the positive results are due to a more faithful implementation of the theory rather than to the teacher variable**.

It remains unclear, however, whether more ordinary physics teachers would be able to teach according to the basis model paradigm. In the future, one should strive for true experimental conditions with teachers drawn randomly and control group

1
2
3 teachers receiving some dummy coaching. **However, it will be difficult to find**
4 **teachers agreeing to take part in a long lasting education program when**
5 **chosen randomly.**
6
7
8
9

10
11 In this study, all intervention group teachers received the same coaching, but it
12 remains unclear why the group of high performing teachers was able to follow the
13 theory more than their colleagues, thus producing bigger effects for their students.
14
15 The one finding for increased importance of continuing professional development
16 seems too general. In fact, the selection of variables or the variables themselves on
17 the level of teachers' subjective beliefs seem to be deficient. More research into
18 teachers' beliefs and professional knowledge is needed.
19
20
21
22
23
24
25
26
27

28 Next to the evaluation of the models of learning processes this study also shows the
29 evaluation of a teacher education program, for which we set up our own model of
30 evaluation as shown in Figure 1.
31
32
33
34

35 One may conclude that an overall education program effect can be seen clearer for
36 classroom actions and students' outcomes than for teachers' beliefs, although the
37 effect sizes are larger for the teachers' beliefs than for student and classroom data.
38
39 Concerning belief changes the intervention group appears as one homogenous unit.
40
41 But classroom and student data both show a similar split-up of the intervention group.
42
43 These two data sources support each other and they appear to be more critical for
44 the students. Evaluations of teacher education programs should therefore include
45 classroom actions and students' outcomes.
46
47
48
49
50
51
52
53
54

55 In addition, our proposed model for the evaluation of this teacher education program,
56 as shown in Figure 1, indicates a direction, in which the education program should
57 work. As a result, a decrease in effect sizes is expected going down the variables.
58
59 This can be confirmed best with teachers/classes from IG High. For the teachers'
60

beliefs, the effects are large with sizes up to $\eta^2 = 0.50$ compared to the comparison group. Effects for classroom actions are also large with sizes up to $d=1.5$. For student outcomes, effects are just mid-sized with sizes for students' perception up to $\eta^2 = 0.05$ and for students' emotion up to $\eta^2 = 0.04$. Students from IG Low do not show effects on this level anymore.

In summary, our proposed model for the evaluation of a teacher education program was able to respond positively to the research questions and performed widely as expected. Our model has shown its usefulness for evaluating a teacher education program.

Overall, this study shows that scaffolding instruction according to the theory of basis models improves physics instruction. It makes sense to distinguish different teaching models according to their teaching aims and to the prerequisites of the students as well as emphasizing all phases of these instructional sequences. Effects of such learning process-oriented **teaching strategy** can be followed through down to student outcomes. Furthermore, the study shows that teachers can actually learn to follow the basis models. Four or five coached lessons in our case proved to be enough for some teachers.

References

- Aebli, H. (1981). *Denken: Das Ordnen des Tuns* [Thinking: Organising the actions]. Band II. Denkprozesse. Stuttgart: Klett-Cotta.
- Aebli, H. (1983). *Zwölf Grundformen des Lehrens* [Twelve basic models of teaching]. Stuttgart: Klett-Cotta.
- Baumert, J., Bos, W., & Watermann, R. (2000). *TIMSS/III: Skalenhandbuch* [Handbook of scales]. Unveröffentlichtes Skript [unpublished handbook] Berlin: Max-Planck-Institute for human development.

- Bransford, J.D. (1984). *The ideal problem solver*. W.H. Freeman: New York.
- Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review*, 31 (1), 21-32.
- Bybee, R. W. (2004). Scientific inquiry and science teaching. In L.B. Flick, & N.G. Lederman (Eds.), *Scientific Inquiry and Nature of Science* (pp. 1-12). Dordrecht (NL): Kluwer Academic.
- Clausen, M. (2004). *Online Rater-Inventar zur Unterrichtsqualität* [Online-rating-tool for lesson quality] (3rd Revision). University of Mannheim.
- Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education – 5th edition*. New York: Routledge Falmer.
- Dewey, J. (1933): *How we think*. Boston: Heath.
- Dewey, J. (1974). *Psychologische Grundfragen der Erziehung* [psychological foundations for education]. München: Ernst Reinhardt .
- Dörner, D., & Wearing, A. (1995). Complex problem solving: Toward a (computer-simulated) theory. In P. A. Frensch, & J. Funke (Eds.), *Complex problem solving: The European Perspective* (pp. 65-99). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Duit, R., & Treagust, D.F. (2003). Conceptual change – A powerful framework for improving science teaching and learning. *International Journal of Science Education*, 25, 671-688.
- Duit, R., & Treagust, D.F. (in press). Conceptual Change: A discussion of theoretical, methodological and practical challenges for science education. *Cultural Studies of Science Education*.
- Fan, X. (2001). Statistical significance and effect size in education research: Two sides of a coin. *The journal of educational research*, 95 (5), 275-282.

- Fischer, H. E., Klemm, K., Leutner, D., Sumfleth, E., Tiemann, R. & Wirth, J. (2005). Framework for Empirical Research on Science Teaching and Learning. *Journal of Science Teacher Education* 16 (4), 309-349.
- Fischer, H. E., Trendel, G., Reyer, Th. & Wackermann, R. (2006). Supporting learning processes – Professional development and in-service training of physics teachers. I. Eilks & B. Ralle (Eds.), *Towards Research-based Science Teacher Education*. Aachen, Shaker, 171-178.
- Fischler, H., & Schröder, H.-J. (2003). Fachdidaktisches Coaching für Lehrende der Physik [physics didactical coaching for teachers of physics]. *Zeitschrift für Didaktik der Naturwissenschaften*, 9, 43-62.
- Fishman, B.J., Marx, R.W., Best, S., & Tal, R.T. (2003). Linking teacher and student learning to improve professional development in systemic reform. *Teaching and Teacher Education*, 19, 643-658.
- Fries, E., Rosenberger, R. (1970). *Forschender Unterricht* [Research-like instruction]. Frankfurt a.M., Germany: Diesterweg.
- Garet, M.S., Porter, A.C., Desimore, L., Birman, B.F., & Yoon, K.S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38(4), 915-945.
- Goodrum, D., Hackling, M., & Rennie, L. (2001). *The status and quality of teaching and learning of science in Australian schools*. Department of Education, Training and Youth Affairs: Canberra.
- Groeben, N., Wahl, D., Schlee, J., & Scheele, B. (1988). *Forschungsprogramm Subjektive Theorien. Eine Einführung in die Psychologie des reflexiven Subjekts* [Research program subjective theories: An introduction to the psychology of reflective minds]. Tübingen, Germany: Francke.

- Guzzetti, B. J., Snyder, T. E., Glass, G., & Gamas, W. S. (1993). Promoting conceptual change in science: A comparative meta-analysis of instructional interventions from reading education and science education. *Reading Research Quarterly*, 28 (2), 116-159.
- Hand, B., & Treagust, D. (1994). Teachers' thoughts about changing to constructivist teaching. *Journal of education for teaching*, 20 (1), 97-112.
- Helmke, A. (2003). *Unterrichtsqualität – erfassen, bewerten, verbessern* [Instructional quality – measurement, evaluation, improvement]. Seelze, Germany: Kallmeyer.
- Helmke, A. & Weinert F.E. (1997). Bedingungsfaktoren schulischer Leistung [Factors influencing student performance]. In F.E. Weinert (Ed.), *Psychologie des Unterrichts und der Schule* (pp. 71-176). Göttingen, Germany: Hogrefe.
- Hofstein, A., & Lunetta, V.N. (1982). The role of laboratory in science teaching. Neglected aspects of research. *Review of Educational Research*, 52, 201-217.
- Hofstein, A., & Lunetta, V.N. (2004). The Laboratory in Science Education: Foundations for the Twenty-First Century. *Science Education*, 88(1), 28-54.
- Joye, B.R., & Weil, M. (1995). *Models of teaching – 5th edition*. Upper Saddle River, NJ: Pearson.
- Klieme, E., Lipowsky, F., Rakoczy, K., & Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht [Quality and efficacy of math instruction]. In M.Prenzel, & L. Allolio-Näcke (Eds.), *Untersuchungen zur Bildungsqualität von Schule, Abschlussbericht des DFG-Schwerpunktprogramms* (pp. 127-146). Münster, Germany: Waxmann.
- Kircher, E., Girwidz, R., & Häußler, P. (2007). *Physikdidaktik* [Didactics of physics]. Braunschweig/Wiesbaden, Germany: Vieweg.

- Kirkpatrick, D. L. (1979). Techniques for evaluating training programs. *Training and Development Journal*, 33(6), 78-92.
- Kirschner, P.A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: an analysis of failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), 75-86.
- Labudde, P., & Duit, R. (2007). Zum Design einer bi-nationalen Videostudie zum Physikunterricht [Design of a bi-national video-study on physics instruction]. In D. Höttecke (Ed.), *Naturwissenschaftlicher Unterricht im internationalen Vergleich* (pp. 631-633). Münster, Germany: Lit.
- Lave, J., & Wenger, E. (2003). *Situated Learning*. Cambridge: Cambridge University Press.
- Lederman, N. G. (2004). Syntax of nature of science within inquiry and science instruction. In L. B. Flick, & N. G. Lederman (Eds.), *Scientific inquiry and nature of science* (pp. 301-317). Dordrecht (NL): Kluwer.
- Lederman, J. S., Lederman, N. G., Wickman, P.-O., & Lager-Nyqvist, L. (2007). *An international, systematic investigation of the relative effects of inquiry and direct instruction*. Presentation at the NARST-conference 2007, New Orleans.
- Lipowsky, F. (2004). Was macht Fortbildungen für Lehrkräfte erfolgreich [What makes teacher education programs successful?]. *Die deutsche Schule*, 96(4), 462-479.
- Luft, J.A. (2001). Changing inquiry practices and beliefs: The impact of an inquiry-based professional development programme on beginning and experienced secondary science teachers. *International Journal of Science Education*, 23(5), 517-534.

- 1
2
3 Lunetta, V. (1998). The school science laboratory: Historical perspectives and
4
5 contexts for contemporary teaching. In Fraser, B., & Tobin, K. (Eds.),
6
7 *International Handbook of Science Education* (pp. 249-264). Dordrecht, the
8
9 Netherlands: Kluwer Academic.
- 10
11
12 Mikelskis-Seifert, S. & Duit, R. (2007). *Physics teacher professional development in*
13
14 *the program "physics in context"*. Unpublished proposal for the NARST
15
16 conference 2008, Baltimore.
- 17
18
19 Osborne, J. F., & Dillon, J. (2008). *Science Education in Europe: Critical Reflections*.
20
21 A report to the Nuffield Foundation. London: King's College.
- 22
23
24 Oser, F., & Baeriswyl, F. (2001). Choreographies of Teaching. Bridging instruction to
25
26 learning. In V. Richardson (Eds.), *AERA's Handbook of Research on Teaching*
27
28 – 4th Edition. Washington: American Educational Research Association.
- 29
30
31 Pauli, C., & Reusser, K. (2003). Teaching-learning scripts in Suisse and German
32
33 math instruction. *Unterrichtswissenschaft*, 31(3), 238-272.
- 34
35
36 Polya, G. (1957). *How to solve it*. New York: Doubleday.
- 37
38
39 Posner, G.J., Strike, K.A., Hewson, P.W., & Gertzog, W.A. (1982). Accomodation of a
40
41 scientific conception: Toward a theory of conceptual change. *Science*
42
43 *Education*, 66(2), 211-227.
- 44
45
46 Reyer, T. (2004). *Oberflächenmerkmale und Tiefenstrukturen im Unterricht* [Surface
47
48 and deep structure in instruction]. Berlin: Logos.
- 49
50
51 Reyer, T.; Trendel, G. & Fischer, H. E. (2004): Was kommt beim Schüler an? –
52
53 Lehrerintentionen und Schülerlernen im Physikunterricht [the effect of teachers'
54
55 intentions and learning processes of students in physics lessons]. In Doll, J. &
56
57 Prenzel, M. (Eds.), *Studien zur Verbesserung der Bildungsqualität von Schule:*
58
59
60

Lehrerprofessionalisierung, Schülerförderung und Unterrichtsentwicklung.

Münster: Waxmann, 195-211.

Rost, D. (2005). *Interpretation und Bewertung pädagogisch-psychologischer Studien*

[Interpretation and evaluation of studies in educational psychology]. Weinheim, Germany: Beltz.

Seidel, T., & Prenzel, M. (2004). Muster unterrichtlicher Aktivitäten im

Physikunterricht [Patterns of classroom actions in physics instruction]. In J. Doll,

& M. Prenzel (Eds.), *Bildungsqualität von Schule: Lehrerprofessionalisierung,*

Unterrichtsentwicklung und Schülerförderung als Strategien der

Qualitätsverbesserung (pp. 177-194). Münster, Germany: Waxmann.

Seidel, T., Prenzel, M., Rimmle, R., Schwindt, K., Kobarg, M., Herweg, C., &

Dalehefte, I.M. (2006). Unterrichtsmuster und ihre Wirkungen. Eine Videostudie

im Physikunterricht [Instructional scripts and their effects. A video study in

physics instruction]. In Prenzel, M., & Allolio-Näcke, L. (Eds.), *Untersuchungen*

zur Bildungsqualität von Schule (pp. 7-19). Münster, Germany: Waxmann.

Schmidkunz, H., & Lindemann, H. (1992). *Das forschend-entwickelnde*

Unterrichtsverfahren [The problem-oriented teaching model]. Magdeburg,

Germany: Westarp Wissenschaften.

Stigler, J., & Hiebert, J. (1997). Understanding and improving mathematics

instruction: An overview of the TIMSS Video Study. *Phi Delta Kappa*, 79(1), 14-

21.

Terhart, E. (2003). Wirkungen von Lehrerbildung: Perspektiven einer an Standards

orientierten Evaluation [Effects of teacher education: Impressions from a

standard-oriented evaluation]. *Journal für LehrerInnenbildung*, 3(3), 8 - 19.

- 1
2
3 TIMSS-Items (1995). Retrieved September 28, 2007, from [http://www.mpib-](http://www.mpib-berlin.mpg.de/TIMSSII-Germany/index.htm)
4
5 [berlin.mpg.de/TIMSSII-Germany/index.htm](http://www.mpib-berlin.mpg.de/TIMSSII-Germany/index.htm).
6
7
8 TIMSS-Items (1999). Retrieved September 28, 2007, from [http://www.mpib-](http://www.mpib-berlin.mpg.de/TIMSSIII-Germany/index.htm)
9
10 [berlin.mpg.de/TIMSSIII-Germany/index.htm](http://www.mpib-berlin.mpg.de/TIMSSIII-Germany/index.htm).
11
12
13 Trendel, G., Wackermann, R., Fischer, H.E. (2007). Lernprozessorientierte
14
15 Fortbildung von Physiklehrern [Learning process oriented in-service training of
16
17 physics teachers]. *Zeitschrift für die Didaktik der Naturwissenschaften (ZfdN)*,
18
19 13, 9-31.
20
21
22 Treagust, D.F. (2007). General Instructional Methods. In S. K. Abell, & N. G.
23
24 Lederman (Eds.), *Handbook of Research on Science Education* (pp. 373-392).
25
26 Mahwah, NJ: Lawrence Earlbaum.
27
28
29 Wagenschein, M. (1970). *Verstehen lernen* [Learning to understand]. Weinheim,
30
31 Germany: Beltz.
32
33
34 Yerrick, R., Parke, H., & Nugent, J. (1997). Struggling to promote deeply rooted
35
36 change: The „Filtering Effect” of teachers’ beliefs on understanding
37
38 transformational views of teaching science. *Science Education*, 81, 137-159.
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1:

Phase	Learning through experience	Concept building	Problem solving
1	Planning of actions	Pre-knowledge	Problem presentation
2	Performance of actions	Introduction of the new concept with an example	Reformulation of problem task
3	First Reflection, construction of meaning	Development of characteristics of the new concept	Development of hypotheses
4	Generalization of experience	Active application of the new concept	Test of hypotheses
5	Abstraction of experience	Application in other contexts	Evaluation of solution

Tab.1: Important learning processes

Table 2:

Oser's original description of phase 2 in <i>basis model</i> learning through discovery	Levels of fulfilment
Performance of actions - Students investigate, experiment, search, categorize...They have an action plan in mind.	0. Students do something, but no action plan is visible.
	1. Students act recipe-like according to a teacher's plan.
	2. Students act rather autonomously within a given action plan.
	3. Students act autonomously and along their own action plan.

Tab. 2: Example of our levels of fulfilment

Table 3:

Variable	Exemplary item	# of items	Reliability(α)	Origin
Teachers' subjective beliefs				
Importance of student process competencies	How important is it for you that students can caing new insights through planned experiments?	6	0.7	Own development
Importance of professional development	I as a teacher want to continually improve professionally.	5	0.8	Own development
Importance of declarative knowledge	How important is it for you as a teacher that your students know exact definitions of physical quantities?	6	0.8	Own development
Classroom actions				
Basis model	Intercoder reliability: $\kappa = 0.8$			Oser
Instructional phase	Intercoder reliability: $\kappa = 0.8$			Oser
Level of fulfilment	Intercoder reliability: $\Gamma = 0.7$			Own addition
Students' perception of instructional quality, student emotions und performance				
Perceived teaching for understanding	How often do you have to explain the reasoning behind in your physics instruction?	5	0.7	TIMSS background questionnaire
Pacing	Our teacher proceeds quickly but does not outrun us.	5	0.9	Instructional quality rating (Clausen)
Perceived structuring aids	Our teacher emphasizes important aspects.	4	0.7	Instructional quality rating (Clausen)
Interest in physics	For activities involving physics I am willing to spend my free time.	4	0.8	TIMSS background questionnaire
TIMSS performance test	Age-appropriate sample of TIMSS-items from population II and III	20	n.a.	TIMSS-tests population II and III

Tab. 3: Selection of variables under investigation at the levels of teachers' beliefs, classroom actions and student outcomes.

Table 4:

	Intervention group	Control group
pre	3.5	3.4
post	3.7	3.2

Tab. 4: Estimated marginal means for teachers' perceived importance of student competencies (range from „1=not important“ to „4=very important“).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 5:

Basis model	Before intervention (n = 18 first videos)	During intervention (n = 62 videos)
Experience learning	41.5 %	29.6 %
Concept building	44.7 %	43.3 %
Problem solving	4.5 %	17.3 %
Other	9.3 %	9.9 %

Tab. 5: Percentage of instruction time per basis model

Table 6:

	High performing teachers	Low performing teachers	control group
pre	3.0	3.5	3.6
post	3.4	3.5	3.5

Tab. 6: Estimated marginal means for teachers' perceived importance of professional development (range from „1=not important“ to „4=very important“).

Figure 1:

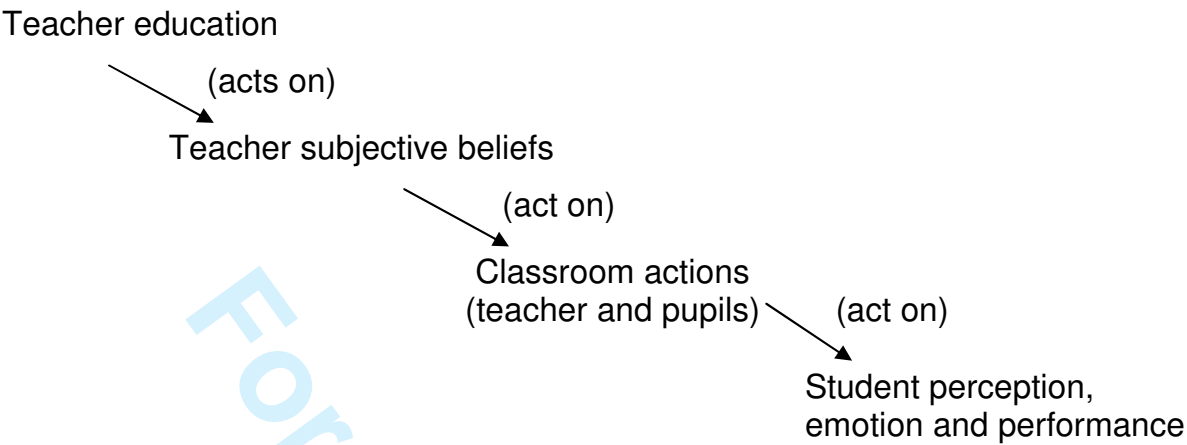


Fig. 1: The model for evaluation of the teacher education program

Figure 2:

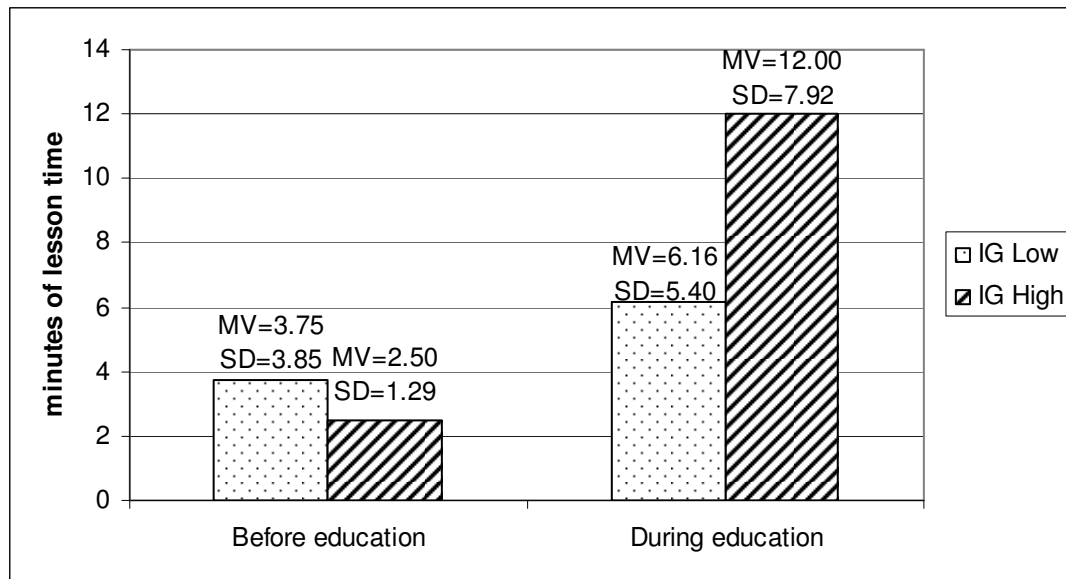


Fig. 2 Average time students participate on level 2 per lesson in comparison between high (IG High) and low (IG Low) performing classes and between prior to education and during the education.

Figure 3:

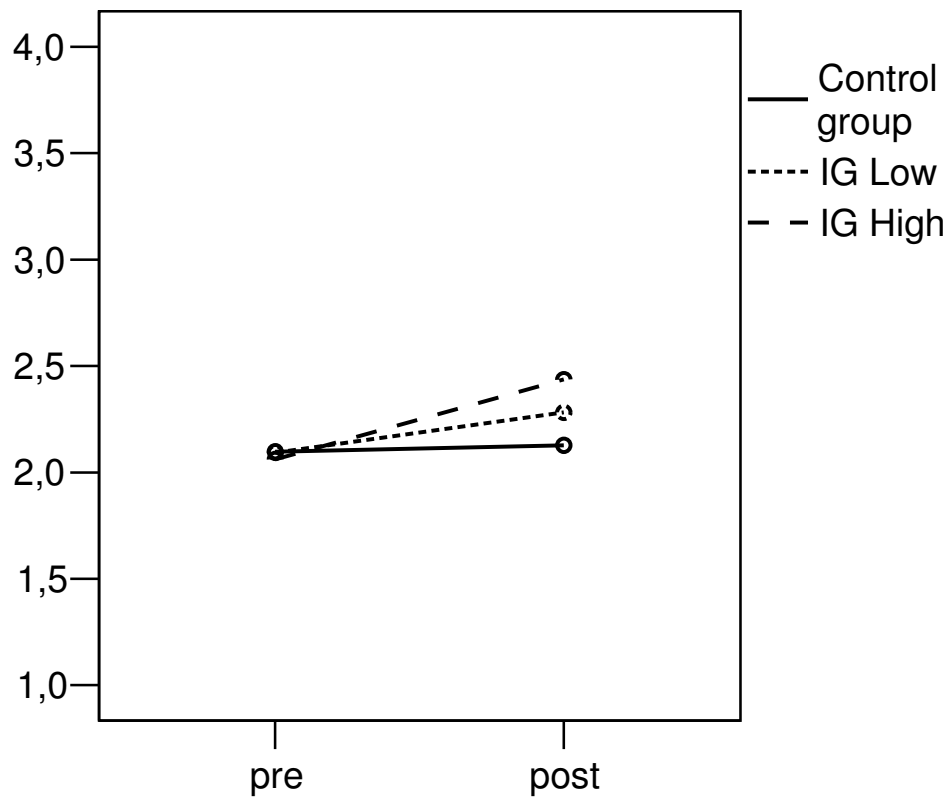


Fig. 3: Estimated marginal means for students' perception of teaching for understanding (range from „1=not at all“ to „4=very much“).

Figure 4:

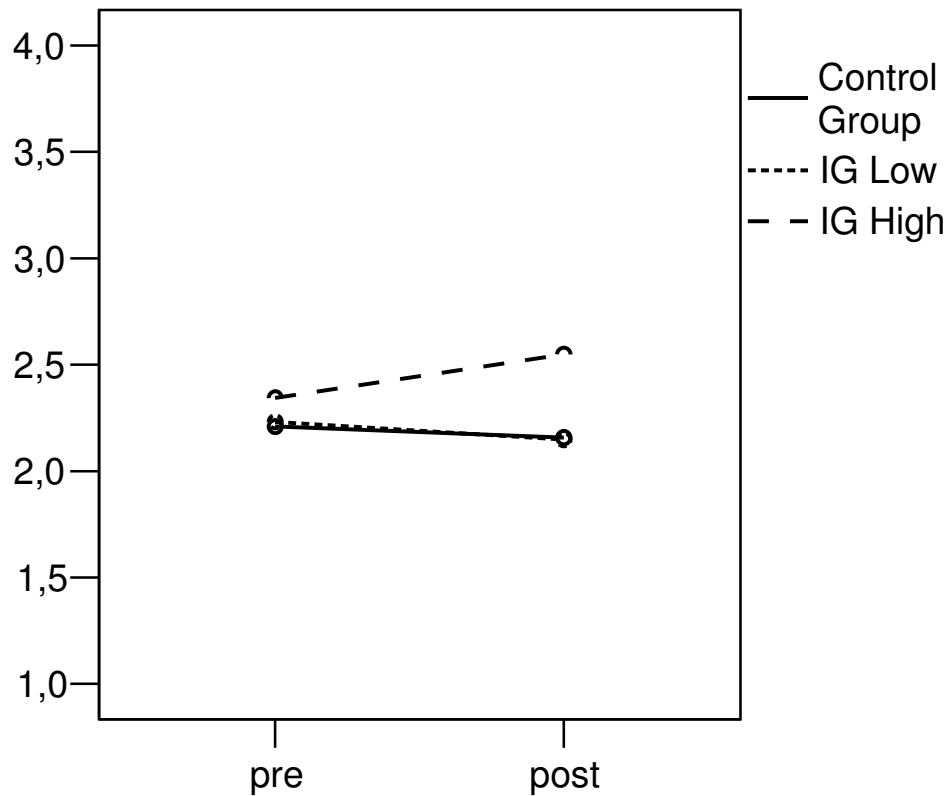


Fig. 4: Estimated marginal means for students' interest in physics (range from „1=not at all“ to „4=very much“).